2006

Technical Characteristics of Two Algebra Progress Monitoring Measures: Reliability, Criterion Validity and Sensitivity to Growth

Serkan Perkmen
Iowa State University

Anne Foegen
Iowa State University, afoegen@iastate.edu

Jeannette R. Olson
Iowa State University, olsonj@iastate.edu

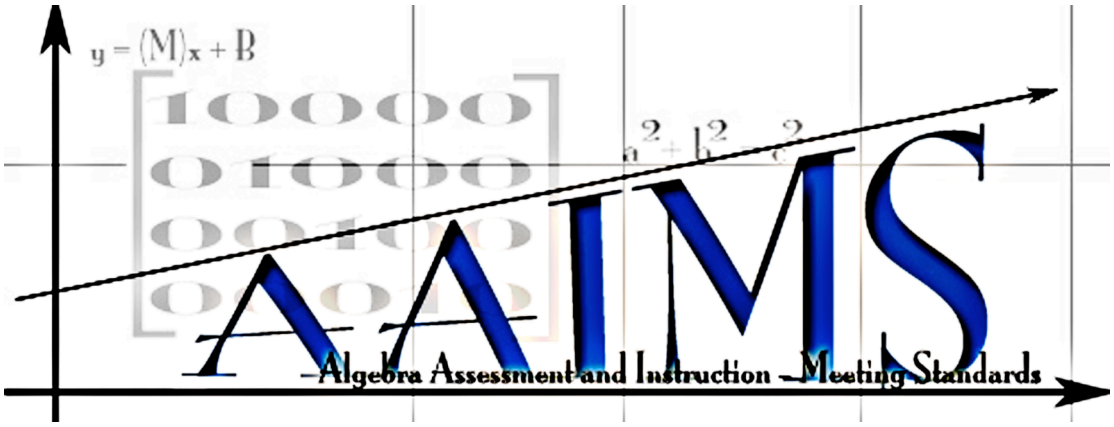
Technical Characteristics of Two Algebra Progress Monitoring Measures: Reliability, Criterion Validity and Sensitivity to Growth

Abstract
The primary purpose of this study was to investigate the reliability, validity of Algebra Foundations and Content Analysis-Multiple probes and to examine the utility of these probes in monitoring student progress over a full school year. In addition, we also examined the use of a third type of probe, Basic Skills, with a small number of students. Our findings revealed that both the Algebra Foundations and Content Analysis-Multiple Choice probes possessed adequate levels of alternate form and test-retest reliability. We examined two types of validity: concurrent and predictive validity. Concurrent validity was assessed by investigating the relationship between probe scores and other indicators of proficiency in algebra including teacher proficiency ratings and standardized test scores. In general, we found probe scores were associated with standardized test scores for students in grades 9-12, but not with teacher ratings of proficiency or standardized test scores for eighth grade students. The predictive validity of the probes was assessed by examining the association between probe scores and other indicators of proficiency including teacher ratings of growth and standardized test scores. Our findings were identical to those for concurrent validity. When examining student progress over time, we found that the Content Analysis-Multiple Choice probes were more sensitive to reflecting student growth than were the Algebra Foundations probes. When investigating student progress over time by class type, we found that only 8th Grade Algebra students showed .5 unit weekly growth on both probes; Algebra 1 students had a mean slope value near this threshold (.47) on the Content Analysis-Multiple Choice probes. This result indicated that the utility of the probes for monitoring student growth may differ for students of different mathematics ability levels.

Disciplines
Curriculum and Instruction | Education | Higher Education | Science and Mathematics Education

Comments
Technical Report #14

PROJECT AAIMS: ALGEBRA ASSESSMENT AND INSTRUCTION – MEETING STANDARDS



Technical Characteristics of Two Algebra Progress Monitoring Measures: Reliability, Criterion Validity and Sensitivity to Growth

Technical Report #14

Serkan Perkmen, M.Ed.

Anne Foegen, Ph.D.

Jeannette Olson, M.S.

Iowa State University

September 2006

PROJECT AAIMS: ALGEBRA ASSESSMENT AND INSTRUCTION – MEETING STANDARDS

Technical Characteristics of Two Algebra Progress Monitoring Measures: Reliability, Criterion Validity, and Sensitivity to Growth:

Technical Report 14

Abstract

The primary purpose of this study was to investigate the reliability, validity of Algebra Foundations and Content Analysis-Multiple probes and to examine the utility of these probes in monitoring student progress over a full school year. In addition, we also examined the use of a third type of probe, Basic Skills, with a small number of students. Our findings revealed that both the Algebra Foundations and Content Analysis-Multiple Choice probes possessed adequate levels of alternate form and test-retest reliability. We examined two types of validity: concurrent and predictive validity. Concurrent validity was assessed by investigating the relationship between probe scores and other indicators of proficiency in algebra including teacher proficiency ratings and standardized test scores. In general, we found probe scores were associated with standardized test scores for students in grades 9-12, but not with teacher ratings of proficiency or standardized test scores for eighth grade students. The predictive validity of the probes was assessed by examining the association between probe scores and other indicators of proficiency including teacher ratings of growth and standardized test scores. Our findings were identical to those for concurrent validity. When examining student progress over time, we found that the Content Analysis-Multiple Choice probes were more sensitive to reflecting student growth than were the Algebra Foundations probes. When investigating student progress over time by class type, we found that only 8th Grade Algebra students showed .5 unit weekly growth on both probes; Algebra 1 students had a mean slope value near this threshold (.47) on the Content Analysis-Multiple Choice probes. This result indicated that the utility of the probes for monitoring student growth may differ for students of different mathematics ability levels.

Full Report

Introduction

Previous work in Project AAIMS has established the reliability and criterion validity of three measures for monitoring student progress in algebra. In Technical Report 10, we reported the technical features of the measures when used for static (i.e., one point in time) measurement of student performance. The three measures (Basic Skills, Algebra Foundations, Content Analysis-Multiple Choice) have acceptable levels of reliability and moderate levels of criterion validity. While it is valuable to have measures that can be used at a single point in time, if teachers want to use the measures to track student progress and inform their instructional decisions, it is important that the measures also reflect changes in student performance over time. In Technical Reports 12 and 13, we presented the results of two semester-long studies of student growth over time on the Basic Skills and Content Analysis-Multiple Choice probes. The study reported here was conducted to replicate the technical adequacy findings of the earlier studies and to examine the degree to which the measures were sensitive to changes in student performance over an entire school year using the Algebra Foundations and Content Analysis-Multiple Choice measures. We also report data for a small sample of students in another district who completed the Basic Skills and Content Analysis-Multiple Choice measures over the course of a school year.

Method

The study described in this report was conducted from September 2005 to April 2006 in District A. This district serves four small towns as well as the rural agricultural areas between the towns. Approximately 7,000 residents reside in the school district. The junior/senior high school has an enrollment of approximately 600 students; about 12 percent of these students receive special education services. Eighteen percent of the district's students are eligible for free and reduced lunch; three percent are of diverse backgrounds in terms of race, culture and ethnicity. During the first data collection session, students completed the algebra criterion measure. All data collection activities involving students were completed during regular class time. Teachers administered all algebra probes.

In addition to the students in District A, we also report growth data for students in District C who participated in a yearlong algebra course. These students were included in the Technical Reports 12 and 13 data analyses for regarding reliability and validity, but we have opted to report their growth data in this document, along with the students in District A who completed the algebra progress monitoring measures over an entire school year. District C is located in a predominantly rural area and serves approximately 17,700 residents in five small towns and a Native American Settlement community. The high school enrolls 488 students in grades 9 through 12. Thirty-nine percent of the districts' students are of diverse backgrounds in terms of race, culture and ethnicity. Approximately 45% of the student population is eligible for free and reduced lunch. Approximately 15% of the student population has been identified as students eligible for special education services.

Participants

Eighty-six students in Districts A and C participated in the study. Two of the students included in Table 1 were enrolled in a special education algebra course. Because of the small

size of this group and the variations in the probes they completed, data for these students were not included in the reliability and validity analyses. Students in the special education algebra class are included in our analysis of growth rates. Likewise, the District C students are only included in our examination of student growth; their data were included in previous technical reports for analyses of reliability and criterion validity. Written parental/guardian consent and written student assent were obtained for all of these students using procedures approved by Iowa State University's Human Subjects Review Committee. Descriptions of the participating students from in district A and district C are provided in Table 1, and Table 2 respectively.

Table 1. Demographic Characteristics of Student Participants by Grade Level for District A

	Total	Grade 8	Grade 9	Grade 10	Grade 12
N	80	18	53	8	1
Gender					
Male	29	8	18	3	0
Female	51	10	35	5	1
Ethnicity					
White	78	18	52	7	1
Black	2	0	1	1	0
Lunch					
Free/Red	11	0	7	4	0
Disability					
IEP	13	0	9	4	0

Table 2. Demographic Characteristics of Student Participants by Grade Level for District C

	Total	Grade 9	Grade 10	Grade 11	Grade 12
N	6	5		1	
Gender					
Male	3	2		1	
Female	3	3		0	
Ethnicity					
White	4	3		1	
Native American	2	2		0	
Lunch					
Free/Red	2	2		0	
Disability					
IEP	2	1		1	

As the data in Table 1 indicate, almost all of the District A participants (over 97%) were white and an average of 66% were in ninth grade, the traditional grade in which students in these districts complete algebra. Fourteen percent participated in federal free or reduced lunch programs and 16% of the participating students were students with disabilities who were receiving special education services. The students participating in the study were drawn from three types of classes. Seventeen students were enrolled in in a PreAlgebra class, 43 in Algebra 1, 18 were 8th grade students taking Algebra 1 and two were in the special education algebra course. Two students who were in the special education course during the fall semester moved to

the PreAlgebra class in the spring semester. In District A, the PreAlgebra course is similar to courses labeled Algebra 1A in Districts B and C.

Additional Information on Students with Disabilities. Because exploring the applicability of the algebra probes to students with disabilities is an important part of Project AAIMS, additional information about the 15 students with disabilities in both districts participating in the project is provided in Table 3.

Table 3. Descriptive Information on the Programs of Students with Disabilities

Characteristic	Quantification
Disability category	13 Entitled Individual (EI) 2 Learning Disability (LD)
% time in general education	Range = 48 - 94%; Mean = 84%
# of students with math goals	6
# of students receiving math instruction in special education classes	2
# of students receiving math instruction in general education classes	13

Students with disabilities earned mean algebra grades of 2.42 [over C+] (range 0.00 [F] to 4.00 [A]). In Districts A and C, the Iowa Tests of Educational Development are used as a district-wide assessment. On average, students with disabilities obtained national percentile rank scores of 43 and 47 in Concepts/Problem Solving and Computation, respectively.

Measures

Algebra Progress Monitoring Measures. Three algebra measures were examined in this study; sample copies of each are provided in the Appendix. The following paragraphs summarize the characteristics of each of the two types of algebra measures used in the study.

Probe A: Basic Skills Measure

Probe A is designed to assess the 'tool skills' that students need to be proficient in algebra. Just as elementary students' proficiency with basic facts is associated with their ease in solving more complex problems, we hypothesize that there are some basic skills in algebra that serve as indicators of overall proficiency. In our discussions with teachers, they frequently commented that many students had difficulty with integers and with applying the distributive property. The items included in the *Basic Skills* measure address solving simple equations, applying the distributive property, working with integers, combining like terms and applying proportional reasoning. The *Basic Skills* probe includes many skills one would assume that students proficient in algebra would be able to complete with reasonable levels of automaticity. Students have five minutes to work on this probe; six parallel forms were used in the study. Each *Basic Skills* probe consists of 60 items; each item is scored as one point if it is answered correctly.

Probe B: Algebra Foundations Measure

The second algebra progress monitoring probe is *the Algebra Foundations* measure. This assessment is designed to reflect five core concepts and skills that were derived from the literature and consultation with experts in mathematics education. The five foundation areas include (1) writing variables and expressions (2) manipulating expressions involving integers,

exponents, and order of operations; (3) graphing expressions and linear equations; (4) solving one-step equations and simplifying equations; and (5) identifying and extending patterns and functions. Our intent with this measure is to assess the extent to which students are proficient at solving problems that address these foundations of early algebra. It is important to note that with this measure, many of the items represent concepts and skills that would be learned as part of pre-algebra or very early instruction in an Algebra I course, if not earlier. We recognize that proficiency on this task is not equivalent to having mastered all the concepts taught in Algebra I, but we hope to determine whether the scores for this measure might serve as an indicator of more general proficiency in algebra. Students have five minutes to work on this probe. The six parallel forms of the *Algebra Foundations* probe consist of 42 items; eight of these require two responses, so 50 total points are possible on this probe. Each student response is scored as one point if it is answered correctly.

Probe E: Content Analysis-Multiple Choice Measure

The *Content Analysis-Multiple Choice* measure consists of 16 items that correspond to the first eight chapters in the textbook that is used in all three districts participating in Project AAIMS. Problems are placed in random order on each probe. Students are directed to circle the correct response from four alternatives and to show their work unless they are confident they can solve the problems mentally. There are six parallel forms for this probe. Students have seven minutes to work on the Content Analysis-Multiple Choice probes.

Scoring for the *Content Analysis-Multiple Choice* probes is done by comparing student responses to a rubric-based key created by the research staff. Each of the 16 problems is worth up to three points. Students earn full credit (three points) by circling the correct answer from among the four alternatives. If students circle an incorrect response and do not show any work, their answer is considered a 'guess;' the total number of guesses is recorded for each probe. In cases where students show work, the scorer compares the student's work to the rubric-based key, and determines whether the student has earned 0, 1, or 2 points of partial credit. The number of points earned across all 16 problems and the number of guesses are recorded and entered in the data files. A final score is computed by subtracting the number of guesses from the total number of points earned on the probe.

Criterion Measures. In order to evaluate the criterion validity of the algebra progress monitoring measures, we gathered data on a variety of other indicators of students' proficiency in algebra. Some of these measures were based on students' performance in class and their teachers' evaluation of their proficiency. Other measures reflected students' performance on standardized assessment instruments.

The classroom-based measures included grade-based measures and teacher ratings. Each student's *algebra grade*, the grade s/he earned in algebra for the yearlong algebra course, was recorded using a four-point scale (i.e., A = 4.0, B = 3.0).

We also included the teachers' evaluations of student proficiency in algebra by asking each teacher to complete a *teacher rating of proficiency* for all the students to whom she/he taught algebra. These ratings were completed at the beginning of the course. Student names were alphabetized across classes to minimize any biases that might be related to particular class sections. Teachers used a 5-point Likert scale (1=low proficiency, 5= high proficiency) to rate each student's proficiency in algebra in comparison to same-grade peers This enabled us to see if

there was a relationship between growth that students showed on both types of probes and the teachers' evaluations of student growth.

Student performance on standardized, norm-referenced assessments was evaluated using school records and with an algebra instrument administered as part of the project. In District A, students complete either the *Iowa Tests of Educational Development* (ITED) or the *Iowa Tests of Basic Skills* (ITBS) each year, depending on their grade level. Students in grades 9 to 12 complete the ITED, while students in grade 8 take the ITBS. District records were used to access students' scores on these instruments; national percentile ranks were used for the analyses. We recorded the Concepts/Problems subtest score (which was identical to the Math Total score) and the Computation subtest score for the ITED and the Math Total score for the ITBS.

Because the district-administered measure did not provide a direct assessment of algebra, so we also administered the *Iowa Algebra Aptitude Test* (IAAT). This norm-referenced instrument is typically used to evaluate the potential of 7th grade students for successful study of algebra in 8th grade. We recorded students' national percentile rank scores on the IAAT and used these data for our criterion validity analyses. Although we recognized the limitations of using this aptitude measure, we were unable to identify a norm-referenced test of algebra achievement. We had some concerns that there might be ceiling effects when using this measure, but these concerns proved to be unwarranted.

Growth Measures. One of the major goals of the AAIMS project is to determine the extent to which the two types of measures reflect student growth over time. We were also interested in exploring whether the growth that students showed on the probes was associated with other indicators of growth. To accomplish these goals, we gathered data using three types of measures reflecting students' growth: *probe slope*, *teacher rating of growth*, and *IAAT gain score.*

The first type of growth measure, which we called *probe slope*, reflects the growth that students showed on both types of probes over the semester. We used ordinary least square regression to calculate each student's slope on each measure. The obtained slope values were calculated to reflect the amount of weekly progress a student demonstrated on a probe type. The second type of measure was the *teacher rating of growth*. At the end of the school year, we asked teachers to rate all the students in their algebra classes. Student names were alphabetized across class periods to minimize any biases that might be related to particular sections. Teachers used a 5-point Likert scale to rate each student's growth in algebra in comparison to same-grade peers. A rating of 1 indicated minimal or no growth, while a rating of 5 represented unusually high growth in comparison to peers. The third type of measure was the *IAAT gain score*, which was calculated by subtracting the total scale raw score for the IAAT form taken at the beginning of the year from the total scale raw score for the IAAT form taken at the end of the year. All students in the project completed Form A of the IAAT at the beginning of the study and Form B at the end. Using correlational analysis, we examined the relationships among these growth variables.

Procedures

Project AAIMS research staff visited each class at the beginning of the school year to present information about the study and gather informed consent. Students completed student assent forms during class and were given parent consent forms to take home. Teachers offered

extra credit to students for returning signed consent forms (regardless of whether parents provided or withheld consent). The research staff also administered the Iowa Algebra Aptitude Test at the beginning and end of the study. Teacher ratings forms were distributed at the beginning (initial teacher rating of student proficiency) and the end (teacher rating of growth) of the study and collected by project staff.

Teachers were instructed to administer two forms of one of the types of probes in the middle of the month and two forms of the other measure at the end of the month. The algebra probes were administered during a portion of each class period. Teachers administered probes according to a schedule with one-week intervals during which they were to give two forms of one type of probe. Some teachers opted to give both probes on the same day; other teachers gave the two probes on two different days. Both participating teachers were able to administer at least one of the two the probes on schedule for the entire year. In District A, general education teachers administered the Algebra Foundations and Content Analysis-Multiple Choice measures; the special education algebra teacher administered Algebra Foundations and Basic Skills. In District C, the general education teacher (of the year long class) administered the Basic Skills and Content Analysis-Multiple Choice measures.

Scoring Reliability

We hired and trained four pre-service teachers (subsequently referred to as "scorers") to score the probes. The hiring process included a demonstration of correct scoring procedures for each type of probe and guided practice activities in which scorers worked with actual student papers. A final activity was the independent scoring of 10 student papers for each of the probe types. We used these probes to evaluate scoring reliability. For each probe, an answer-by-answer comparison was conducted and an interscorer reliability estimate was calculated by dividing the number of agreements by the total number of answers scored. These individual probe agreement percentages were then averaged across all the selected probes of a common type to determine an overall average. After training, the scorers' mean interscorer agreement rates were 97.9% for the Algebra Foundations probes (range = 97.1% to 99.1%) and 95.84% for the Content Analysis-Multiple Choice probes (range = 94.27% to 96.90%). The mean interscorer agreement rate for the Basic Skills probes was 98.95% (range = 98.45% to 99.63%). Scorers were informed that we would be checking their scoring accuracy levels throughout the project; they were able to earn bonus pay for maintaining high levels (i.e., >96% agreement) of accuracy in their scoring.

Following training, each scorer was assigned five classes with two forms of a probe per class to score (a total of 10 class sets of probes twice each month). Readers should note that the total of 20 classes includes additional algebra classes in Project AAIMS whose data are reported in other technical reports. Scorers also completed the data entry for the classes they were scoring. For each scorer, we conducted a scoring reliability on two of the ten class sets in each scoring period (i.e., twice each month) by re-scoring all of the probes in those sets. The results of these interscorer reliability analyses are reported in the following section.

Results

Scoring Reliability

Interscorer agreement rates revealed that scorers had high reliability on both types of probes. A total of 112 interscorer reliability checks were conducted across the four scorers

throughout the 2005-06 school year. The range of agreement for Algebra Foundations probes was between 96.1% and 100% with a mean of 97.9%. For Content Analysis-Multiple Choice probes, the interscorer agreement rates ranged from 94.5% to 100%, with a mean of 98.7%. For the Basic Skills probes, the range of agreement was between 98.1% and 100%, with a mean of 99.1%.

Descriptive Data on Score Ranges and Distributions

Table 4 lists the ranges, means, and standard deviations for the Algebra Foundations probes. For these probes, the number of correct answers was recorded. The total number of points possible for this probe was 50.

A close examination of Table 4 reveals two important points. First, although there were fluctuations in scores on these probes over the year, probe scores generally increased as the year went on. This finding suggests that students were improving their proficiency in completing the types of problems on these probes. Second, the standard deviations were substantial (one-fourth to one-third of the magnitude of the means), suggesting that the measures would be helpful in distributing students based on scores obtained on both probes. This finding is especially important if the probe data are to be used to identify students who are especially strong or weak in algebra.

Table 4. Descriptive Data for Algebra Foundations Probes Across Administration Sessions – Raw Scores

Time Period	Probe	N	Score	Range	Mean	Standard Deviation
Mid-September	B-31	74	5 min.	5 - 28	19.41	5.34
	B-32	73	5 min.	10 – 41	24.11	6.75
Mid-October	B-33	75	5 min.	5 – 33	17.75	6.46
	B-34	75	5 min.	7 – 37	23.16	7.69
Mid-November	B-35	71	5 min.	5 – 37	21.35	7.23
	B-36	69	5 min.	7 – 36	22.12	6.25
Mid-December	B-31	72	5 min.	9 – 42	23.79	7.02
	B-32	55	5 min.	10 - 44	27.44	8.82
End-January	B-33	71	5 min.	10 - 45	25.92	9.19
	B-34	71	5 min.	12 - 47	29.52	9.30
End-February	B-35	72	5 min.	12 - 43	27.31	7.56
	B-36	72	5 min.	7 - 43	26.15	7.93
Mid-April	B-31	69	5 min.	11 - 45	27.17	8.65
	B-32	69	5 min.	9 - 50	30.48	10.43

On the *Content Analysis-Multiple Choice* probes, the correct score represents the number of points earned on the probe (each of the 16 problems was worth up to 3 points) and the guess score represents the number of guess responses. The total possible correct and guess scores were 48 and 16, respectively. Table 5 presents the ranges, means, and standard deviations for the *Content Analysis-Multiple Choice* probes. A close examination of this table reveals the same patterns as in the data for the Algebra Foundations probes. First, correct scores seemed to increase gradually as the year went on. Second, the standard deviations were substantial (about

one-third of the magnitude of the means), suggesting that the measures would be helpful in spreading out students on the basis of scores received on these probes. This finding can be used as evidence that like Algebra Foundations probes, Content Analysis-Multiple Choice probes would be helpful in identifying students who are especially strong or weak in algebra.

Table 5. Descriptive Data for Content Analysis-Multiple Choice Probes Across Administration Sessions – Raw Scores

End-September	E-31	73	Correct	4 - 31	14.40	5.23
	E-31	73	Guess	0 - 7	1.44	1.50
	E-32	72	Correct	3 - 30	16.88	5.42
	E-32	72	Guess	0 - 11	1.81	1.87
End-October	E-33	72	Correct	8 - 39	19.53	5.73
	E-33	72	Guess	0 - 9	1.19	1.78
	E-34	70	Correct	3 - 39	16.57	6.54
	E-34	70	Guess	0 - 11	1.41	1.94
End-November	E-35	70	Correct	4 - 42	19.79	7.37
	E-35	70	Guess	0 - 12	1.57	2.34
	E-36	72	Correct	6 - 42	22.64	7.49
	E-36	72	Guess	0 - 12	1.61	2.22
Mid-January	E-31	72	Correct	7 - 46	24.72	8.40
	E-31	72	Guess	0 - 11	1.67	2.38
	E-32	72	Correct	9 - 48	26.82	10.23
	E-32	72	Guess	0 - 12	1.49	2.31
Mid-February	E-33	70	Correct	9 - 43	23.91	8.93
	E-33	70	Guess	0 - 13	1.31	2.11
	E-34	69	Correct	5 - 48	24.36	9.91
	E-34	69	Guess	0 - 11	1.49	2.29
End-March	E-35	69	Correct	7 - 46	24.30	9.70
	E-35	69	Guess	0 - 8	1.25	1.31
	E-36	69	Correct	7 - 45	26.90	9.40
	E-36	69	Guess	0 - 6	1.03	1.20
End-April	E-31	71	Correct	9 - 48	27.94	9.90
	E-31	71	Guess	0 - 7	1.11	1.65
	E-32	71	Correct	9 – 48	29.14	9.81
	E-32	71	Guess	0 – 11	1.27	2.26

We also examined whether scores obtained on the Algebra Foundations and Content Analysis-Multiple Choice probes differed by class type. As discussed earlier, students in three types of classes were participating in the study: Algebra 1 students, 8th Grade Algebra students, and PreAlgebra students.

With regard to the probe means, we put forth two hypotheses for each of the probe types. The first was that 8th Grade Algebra students would have the highest level of performance in every period followed by Algebra 1 students. The second was that mean scores for all the class types would increase as the semester went on. The means and standard deviations by class type for the Algebra Foundations probes are reported in Table 6; those for the Content Analysis-

Multiple Choice probes are in Table 7. A close examination of Table 6 reveals that the first hypothesis was fully supported since 8th Grade Algebra students showed the highest level of performance in each period followed by Algebra 1 students; however we failed to provide evidence that the second was fully supported for all the class types because the scores on these probes did not seem to increase gradually as the year went on. The data in Table 6 indicate a similar pattern in the first three administration periods because mean scores for each class type did not seem to increase in these periods. When we moved from the third period (Mid-November) to the forth period (Mid-December), mean scores substantially increased for all class types. When we examined mean scores for the rest of the year (the last four periods), we found two differences between class types in terms of growth. PreAlgebra students' mean scores did not seem to increase as the rest of the year went on. 8th Grade and Algebra 1 students' mean scores seemed to increase gradually except for the interval from End-January to End-February periods.

We examined the same two hypotheses with regard to the Content Analysis-Multiple Choice probes. As Table 7 indicates, we fully supported the first hypothesis that 8th Grade Algebra students received the highest mean scores in every period followed by Algebra 1 students. The second hypothesis that means scores for each class type would gradually increase is fully supported for only PreAlgebra students. Students in Algebra 1 and 8th Grade Algebra, however, had means scores that declined from January to February, but then increased in each following month. Although we did not find a consistent pattern of increases from one administration to the next, the general pattern was one of increase across the school year. Specific analyses examining the growth characteristics of the measures are reported in a later section.

Table 6. Descriptive Data for Algebra Foundations Probes by Class Type

Time Period/Class Type	N	Range	Mean	Standard Deviation
Mid-September				
Algebra 1A	13	8 - 28	15.85	5.55
Algebra 1	43	14 - 34	22.35	5.23
8th Grade Algebra	18	15 - 32	24.33	4.11
Mid-October				
Algebra 1A	15	6.5 - 28.5	14.30	6.46
Algebra 1	42	10 - 33.5	21.03	6.19
8th Grade Algebra	18	16 - 34	24.22	4.98
Mid-November				
Algebra 1A	13	5 - 27	15.38	5.87
Algebra 1	41	6 - 37	22.45	6.24
8th Grade Algebra	17	17 - 33	24.09	4.38
Mid-December				
Algebra 1A	13	11.5 - 32	19.07	5.88
Algebra 1	41	10 - 42.5	25.47	7.37
8th Grade Algebra	18	22 - 38.5	29.83	4.05
End-January				
Algebra 1A	14	11.5 - 35.5	18.86	6.80
Algebra 1	41	11.5 - 46	28.08	8.45
8th Grade Algebra	16	24.5 - 44	34.53	5.28
End-February				
Algebra 1A	14	9.5 - 34	19.71	6.38
Algebra 1	41	14 - 43	26.43	6.45
8th Grade Algebra	18	24 - 42	32.58	5.29
Mid-April				
Algebra 1A	13	13 - 27	19.54	4.94
Algebra 1	38	10 - 48	29.45	8.88
8th Grade Algebra	18	27 - 45	35.92	6.83

Table 7. Descriptive Data for Content Analysis-Multiple Choice Probes by Class Type

Time Period/Class Type	N	Range	Mean	Standard Deviation
End-September				
Algebra 1A	13	1-14.5	8.50	3.27
Algebra 1	43	7-28	14.81	4.38
8th Grade Algebra	17	11-23	16.29	3.17
End-October				
Algebra 1A	13	4-21	11.38	4.90
Algebra 1	41	6-39	17.78	6.24
8th Grade Algebra	18	9-30	18.61	4.45
End-November				
Algebra 1A	12	1-27	11.67	8.08
Algebra 1	42	7-42	20.50	7.28
8th Grade Algebra	18	12-31	23.00	5.40
Mid-January				
Algebra 1A	15	2.5-28.5	12.90	6.92
Algebra 1	39	11-40.5	24.60	7.85
8th Grade Algebra	18	22.5-47	32.80	6.49
Mid-February				
Algebra 1A	16	2-34	14.06	7.75
Algebra 1	39	11-42	23.29	8.00
8th Grade Algebra	16	17-46	30.53	8.75
End-March				
Algebra 1A	15	5.5-26	15.43	5.85
Algebra 1	40	7.5-45	24.98	9.13
8th Grade Algebra	14	21.5-44.5	32.67	6.37
End-April				
Algebra 1A	14	7-39	17.86	8.67
Algebra 1	39	9-43	26.79	9.29
8th Grade Algebra	18	22-47	35.94	6.60

Reliability of Probe Scores

The alternate form reliability of individual probes was evaluated by examining the correlation between two forms of a probe given during the same data collection session. The results of these analyses are reported in Table 8. Alternate form reliability estimates for the Algebra Foundations probes ranged from .72 to .91, with five of the seven correlations exceeding .80. For the Content Analysis-Multiple Choice measure, the estimates ranged from .48 to .91. The particularly low reliability estimate occurred during the first administration and may be related to students' limited knowledge of the Algebra 1 content included in this measure. During the second half of the school year, the reliability coefficients were consistently above .80. In general, we did not find substantial differences in the alternate form reliability of these two measures.

We hypothesized that as the semester went on and students became more familiar with the probes, alternate form reliabilities would increase. Although the last alternate form reliability

score (.84) was found to be higher than the beginning one (.77) for the Algebra Foundations probes, we were not able to fully support this hypothesis for these probes because there were substantial fluctuations in reliability scores during the year. The reliability coefficients for the Content Analysis-Multiple Choice probes increased gradually from month to month except from January to February.

Table 8. Alternate Form Reliability Results for Single Probes

Time Period	Probes	Reliability
Algebra Foundations		
Mid-September	A-31 and A-32	.77
Mid-October	A-33 and A-34	.85
Mid-November	A-35 and A-36	.72
Mid-December	A-31 and A-32	.86
End-January	A-33 and A-34	.91
End-February	A-35 and A-36	.84
Mid-April	A-31 and A-32	.84
Content Analysis-Multiple Choice		
End-September	E-31 and E-32	.48
End-October	E-33 and E-34	.69
End-November	E-35 and E-36	.78
Mid-January	E-31 and E-32	.83
Mid-February	E-33 and E-34	.80
End-March	E-35 and E-36	.88
End-April	E-31 and E-32	.91

Note: All correlations were significant at $p < .01$.

We assessed the test retest reliability of the probes by examining the correlation between the mean of two forms of a probe administered across two data collection time periods. For example, the two scores on the Algebra Foundations probes administered in mid-September were averaged and then correlated with the mean of the two scores on the Algebra Foundations probes administered in mid-October. Readers should note that this four-week time interval is much longer than what is typically used for evaluations of test-retest reliability. Table 9 presents the results of the test-retest reliability analyses. In general, we found that the Algebra Foundations probes possessed higher test-retest reliabilities (ranging from .80 to .91) than the Content Analysis-Multiple Choice probes (ranging from .64 to .88).

As with alternate form reliability, we hypothesized that as the semester went on, test-retest reliability would also increase. Although the reliability score obtained in the last administration is higher than the one in the first administration for both types of probes, there were some non-linear trends or fluctuations in the reliability coefficients over the year. Regarding the reliability of Algebra Foundations probes, reliability coefficients did not increase in the first three periods, then demonstrated incremental increases with each subsequent administration. For the Content Analysis-Multiple Choice probes, reliability estimates showed fluctuations over the year. For example, the reliability coefficient decreased when November and January scores were correlated, then rebounded to an even higher level when scores from Mid-January and Mid-February period were correlated.

Table 9: Test-Retest Reliability Results for Aggregated Probes

Time Period	Reliability
Algebra Foundations	
Mid-September and Mid-October	.81
Mid-October and Mid-November	.81
Mid-November and Mid-December	.80
Mid-December and End-January	.87
End-January and End-February	.89
End-February and Mid-April	.91
Content Analysis-Multiple Choice	
End-September and End-October	.64
End-October and End-November	.80
End-November and Mid-January	.77
Mid-January and Mid-February	.85
Mid-February and End-March	.82
End-March and End-April	.88

Note: All correlations were significant at $p < .01$.

Concurrent Validity

We examined the concurrent validity of the measures by correlating scores on the probes with the criterion measures that served as additional indicators of students' proficiency in algebra. The indicators we used included teachers' evaluations of student proficiency and scores obtained from norm-referenced tests: the Iowa Algebra Aptitude Test (IAAT), ITED-Computation, ITED-Concept and ITBS scores. The results of the criterion validity analyses are presented in Table 10. Reader should note that because the alternate form reliability of the Content Analysis-Multiple Choice probes was extremely low for the initial (end-September) period, we used both end-September and end-October scores to assess the concurrent validity of Content Analysis-Multiple Choice probes. Including both sets of data allows readers to better evaluate the effects of reliability differences on the concurrent validity of the probes.

Table 10 presents the concurrent validity results. We correlated students' scores on the September and October probes with the fall teacher rating of student proficiency and the IAAT administered at the beginning of the study (labeled "Pre-IAAT" in Table 10). We also correlated students' scores at the end of the study (mid-April and mid-January) with their scores on the IAAT administered at the end of the study ("Post-IAAT"), ITED-Computation, ITED Concept and ITBS scores. With the exception of the nonsignificant correlations between probe scores and ITBS scores, the correlation coefficients ranged from .45 to .79 (see Table 10). In general, the coefficients for the Content Analysis-Multiple Choice measure were comparable to or higher than those for the Algebra Foundations measure, suggesting slightly stronger criterion validity for the Content Analysis-Multiple Choice probes. The data in Table 10 indicate that scores on both the Algebra Foundations and the Content Analysis-Multiple Choice probes were moderately or highly correlated with criterion measures, which supported their concurrent validity.

Table 10. Concurrent Validity Results

Time Period	Teacher Ratings			Pre-IAAT			Post-IAAT			ITED-Computation			ITED-Concept			ITBS		
Algebra Foundations	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*
Mid-September	13	.61	.03	71	.73	<.01												
Mid-April							62	.57	<.01	52	.74	<.01	52	.56	<.01	17	.17	.51
Content Analysis-Multiple Choice																		
End-September	13	.45	.13	70	.66	<.01												
End-October	13	.72	<.01	69	.62	<.01												
End-April							64	.75	<.01	54	.79	<.01	54	.62	<.01	17	.30	.24

We also examined concurrent validity to see if the magnitude of the correlation coefficients differed by class type. These results are reported in Table 11. Readers are reminded that teacher ratings were only available for PreAlgebra students, so the sample sizes for those cells are small. When we looked at the correlations between beginning of the year probe scores (mid-September, end-September, and end-October) and criterion measures including fall teacher ratings and Pre-IAAT, we found different results for each class type. Across both the beginning and the end of the academic year, none of the correlations for 8th Grade Algebra students were statistically significant. We found that the probe scores from the beginning of the school year had moderate or high correlations with teacher ratings and Pre-IAAT for PreAlgebra students. For Algebra 1 students, the earliest scores were moderately correlated with Pre-IAAT scores. In general, the coefficients for both types of probes were similar to or stronger for PreAlgebra students as compared to Algebra 1 students.

When we examined the correlations between the probe scores from the end of the school year (mid-April or end-April) and concurrent criterion measures including the Post-IAAT, ITED-Computation, ITED-Concept, and ITBS scores, we again found different results for each class type. For PreAlgebra students, Content Analysis-Multiple Choice probe scores were significantly correlated with the criterion measures, while the Algebra Foundations scores were not. For Algebra 1 students, the correlations between Content Analysis-Multiple Choice probe scores and the criterion measures were higher than the correlations between the Algebra Foundations probes and the criterion measures.

Predictive Validity

We examined the predictive validity of the measures by correlating scores on the probes completed at the beginning of the school year with students' scores on criterion measures from the latter part of the school year that served as additional indicators of students' proficiency in algebra. The indicators we used included Post IAAT scores, teacher ratings of growth, algebra grades, ITED scores, and ITBS scores. The results, presented in Table12, indicate that both probes were significantly correlated with Post-IAAT scores, algebra grades, and ITED scores. Correlation coefficients ranged from .39 to .65. However, both probes were not significantly correlated with other indicators including teacher growth ratings, and ITBS scores. The nonsignificant relationship between probe scores and teacher growth ratings may stem from the fact that there was small variability in teacher growth rating scores. Almost 75% of the students were rated as 3

We also examined predictive validity to see if the correlation coefficients differed by class type. These results are presented in Table 13. As with the earlier analyses by class type, readers are reminded to interpret the results with caution given the small sample sizes for Pre-Algebra and 8th Grade Algebra students. As with the concurrent validity analyses, we included both end-September scores and end-October scores for the Content Analysis- Multiple Choice probes in our analyses by class type.

Table 11. Concurrent Validity by Class Type

	Teacher Ratings			Pre-IAAT			Post-IAAT			ITED-Computation			ITED-Concept			ITBS		
Algebra Foundations	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*
Mid-September																		
PreAlgebra	13	.61	.03	13	.94	<.01												
Algebra 1	-	-	-	41	.58	<.01												
8th Grade Alg.	-	-	-	17	.32	.21												
Mid-April																		
PreAlgebra							13	.32	.29	13	.34	.26	13	-.03	.91	-	-	-
Algebra 1							31	.52	<.01	38	.71	<.01	38	.46	<.01			
8th Grade Alg.							18	-.32	.19	-	-	-	-	-	-	17	.17	.52
Content Analysis-Multiple Choice																		
End-September																		
PreAlgebra	13	.45	.13	13	.69	.01												
Algebra 1	-	-	-	41	.53	<.01												
8th Grade Alg.	-	-	-	16	-.26	.33												
End-October																		
PreAlgebra	13	.72	<.01	13	.54	.06												
Algebra 1	-	-	-	39	.58	<.01												
8th Grade Alg.	-	-	-	17	-.07	.80												
End-April																		
PreAlgebra							14	.89	<.01	14	.71	<.01	14	.59	.03	-	-	-
Algebra 1							32	.70	<.01	39	.76	<.01	39	.51	<.01	-	-	-
8th Grade Alg.							18	-.10	.68	-	-	-	-	-	-	17	.31	.23

Table 12. Predictive Validity Results for All Students

Time Period/ Class Type	Post-IATT			Teacher Growth Rating			Algebra Grade			ITED Comp			ITED Concept			ITBS		
Algebra Foundations	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*	N	*r*	*p*
Mid-September	66	.61	< .01	74	.05	.63	74	.39	< .01	56	.65	< .01	56	.54	< .01	17	.15	.57
Content Analysis- Multiple Choice																		
End-September	65	.62	< .01	73	.16	.16	73	.54	< .01	56	.62	< .01	56	.51	< .01	16	-.18	.49
End-October	65	.57	< .01	72	.06	.58	72	.51	< .01	54	.64	< .01	54	.56	< .01	17	-.09	.72

The predictive validity pattern for the 8th Grade Algebra students was similar to the concurrent validity results, with only a small number of significant coefficients. All of these relations were between the Content Analysis-Multiple Choice measure and teacher growth ratings or algebra grades. For the Algebra Foundations measure, the correlations were stronger for PreAlgebra students with Post-IAAT, and ITED-Concept scores. The Algebra 1 students had stronger relations between the Algebra Foundations measure and algebra grades and ITED-Computation. For the Content Analysis-Multiple Choice September data, relations with the Post-IAAT and algebra grades were stronger for PreAlgebra students, while relations with the ITED subtests were stronger for the Algebra 1 students (and non-significant for PreAlgebra students). The October Content Analysis-Multiple Choice data were not significantly related to the any of the criterion measures for PreAlgebra students. For Algebra 1 students, the relations were similar, but slightly stronger than the results obtained with the September data.

Growth

Readers are reminded that in this section of the paper, we include three groups of students: the students enrolled in general education courses in District A, as well as the students in Special Education Algebra in District A and the students in District C enrolled in a year-long algebra course. Students in the District A Special Education Algebra course completed Algebra Basic Skills and Algebra Foundations probes during the year. Students in the District C yearlong course completed Basic Skills and Content Analysis-Multiple Choice probes. We completed three types of analyses of growth. First, we plotted mean scores over time to visually examine the data. Second, we computed individual students' slope values and determined mean slopes for each type of measure. Finally, we examined correlations between the slopes and other indicators of growth.

We first examined growth by plotting mean scores to visually examine the data. We hypothesized that students' scores on the measures would increase over the course of the year as students received algebra instruction. Figure 1 shows the growth that each type of general education class in District A demonstrated on the Algebra Foundations probes over the year. While the general trend across the entire year was one of growth, students in these classes demonstrated periods during which their scores held constant. As we would have predicted, 8th Grade Algebra students showed the most growth and obtained the highest scores in each administration period, while PreAlgebra students had the lowest scores and demonstrated slower growth. Figure 2 shows the performance of the two students in Special Education Algebra on the Algebra Foundations measure. For one student, there was a general pattern of increase, with a high degree of individual variability. For the second student, there appeared to be little change in performance levels on this measure over the course of the year.

Table 13. Predictive Validity by Class Type

Time Period/ Class Type	Post-IATT			Teacher Growth Rating			Algebra Grade			ITED Comp			ITED Concept			ITBS		
Algebra Foundations	n	*R*	*p*	N	*R*	*p*	n	*r*	*p*	n	*r*	*p*	n	*r*	*p*	*n*	*r*	*P*
Mid-September																		
Algebra 1A	13	.81	<.01	13	-.29	.32	13	.19	.53	13	.52	.07	13	.64	.02	-	-	-
Algebra 1	35	.48	<.01	43	.05	.76	43	.38	.02	42	.58	<.01	42	.36	.02	-	-	-
8th Grade Algebra	18	.00	.98	18	.28	.26	18	.07	.77	-	-	-	-	-	-	17	.14	.57
Content Analysis- Multiple Choice																		
End-September																		
Algebra 1A	13	.78	<.01	13	-.04	.89	13	.60	.03	13	.42	.16	13	.27	.39	-	-	-
Algebra 1	35	.52	<.01	43	.16	.30	43	.55	<.01	42	.54	<.01	42	.41	<.01	-	-	-
8th Grade Algebra	17	.02	.91	17	.42	.09	17	.46	.06	-	-	-	-	-	-	16	-.18	.50
End-October																		
Algebra 1A	13	.43	.15	13	-.49	.08	13	.19	.53	13	.41	.16	13	.52	.06	-	-	-
Algebra 1	34	.60	<.01	41	.26	.10	41	.68	<.01	40	.59	<.01	40	.45	<.01	-	-	-
8th Grade Algebra	18	-.04	18	18	.61	<.01	18	.14	.57	-	-	-	-	-	-	17	-.09	.72

Figure 1. Mean Scores on Algebra Foundations Probes for District A General Education Classes

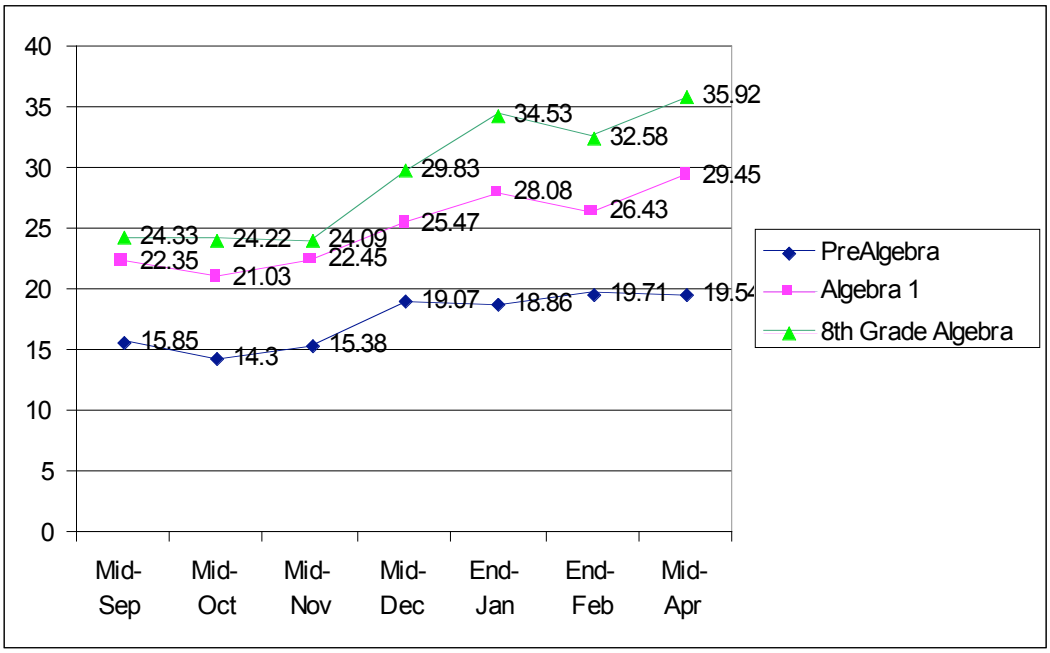
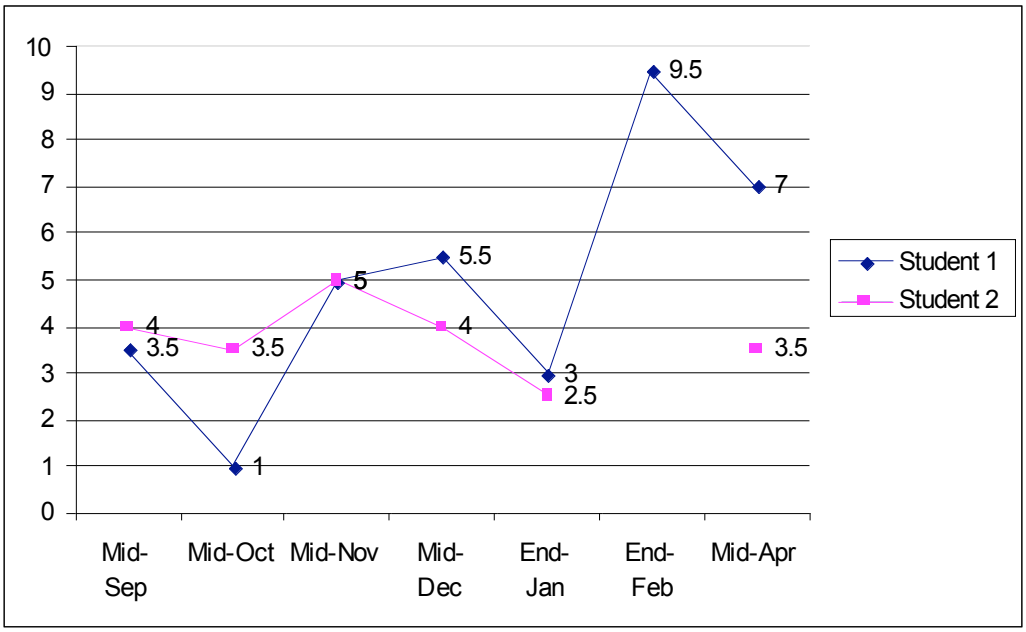


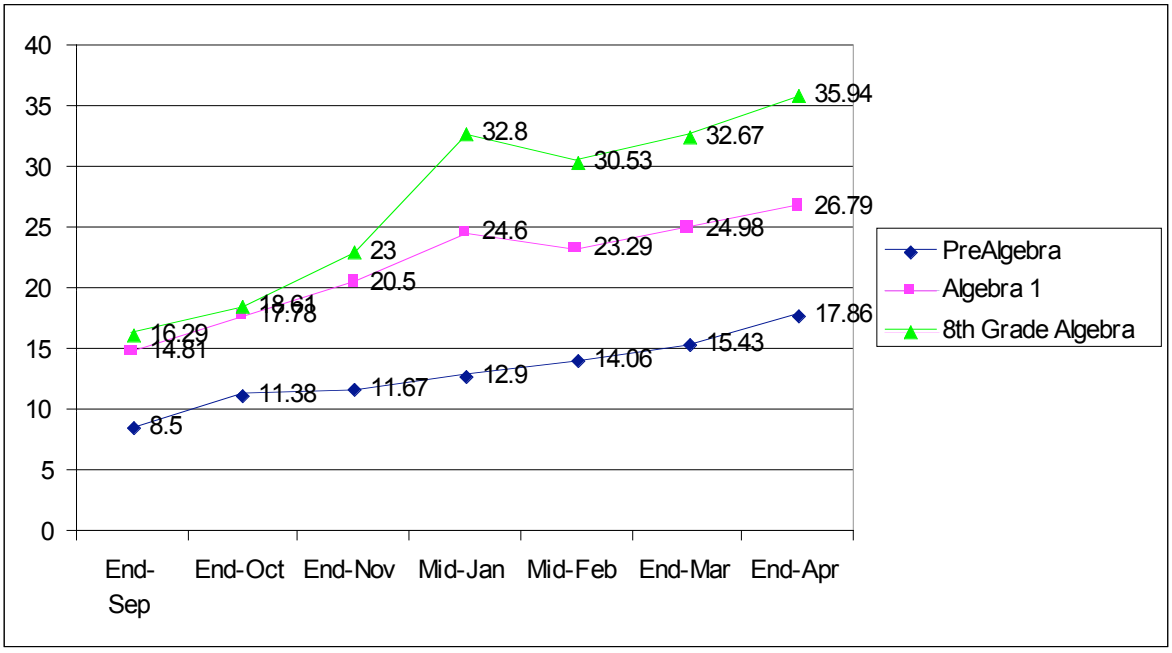
Figure 2. Individual Scores on Algebra Foundations Probes for each Special Education Algebra Student



We conducted the same analyses for the Content Analysis-Multiple Choice probes. The results for general education classes in District A, presented in Figure 3, revealed more consistent and dramatic growth on the Content Analysis-Multiple Choice measure than on the Algebra Foundations measure. Figure 4 shows the growth that students in District C

demonstrated on the same measure. As with the results for District A, students in the yearlong algebra class in District C demonstrated substantially more growth on the Content Analysis-Multiple Choice probes than on the Algebra Foundations measure. It is important to note that across all class types, the mean scores for the last administration of the school year were almost twice as high as the mean scores for the first administration. This finding suggests that the Content Analysis-Multiple Choice probes are better indicators of student growth than are the Algebra Foundations Probes.

Figure 3. Mean Scores on the Content Analysis-Multiple Choice Probes for General Education Classes in District A



We next examined the mean scores for the Basic Skills probes. Figure 5 shows the mean scores for students in District C over the year. The figure suggests that performance on the Basic skills measure did not increase much for this group of students over the school year. Figure 6 reveals that while one student's scores on the Basic Skills measure were relatively constant, the other student's performance declined over the school year.

Figure 4. Mean Scores on the Content Analysis -Multiple Choice Probes in District C

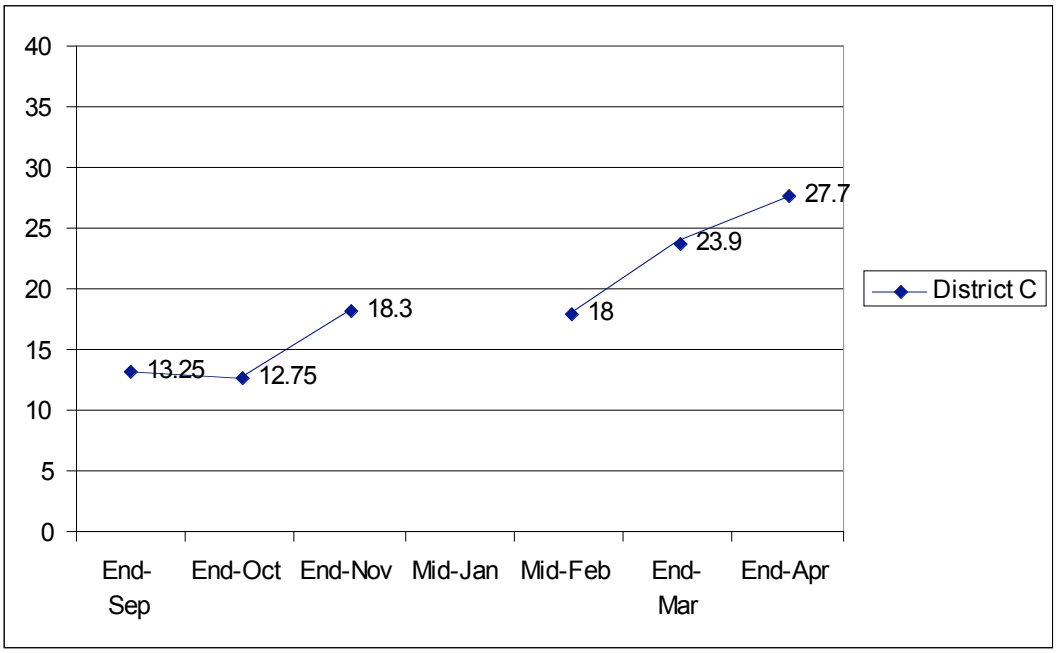


Figure 5: Mean Scores on Basic Skills Probes in District C

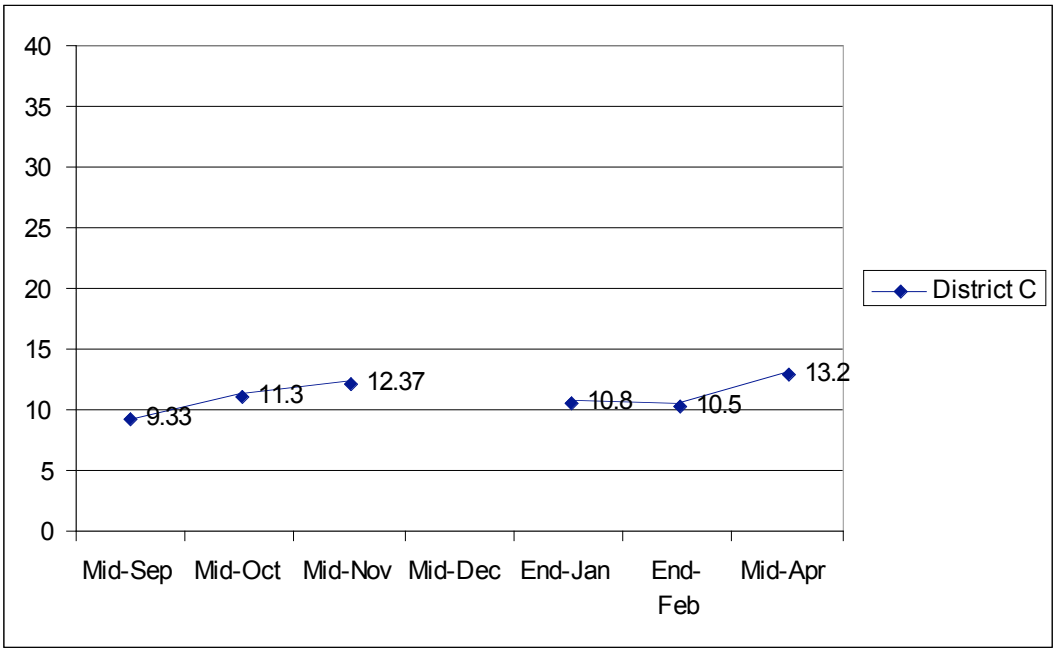
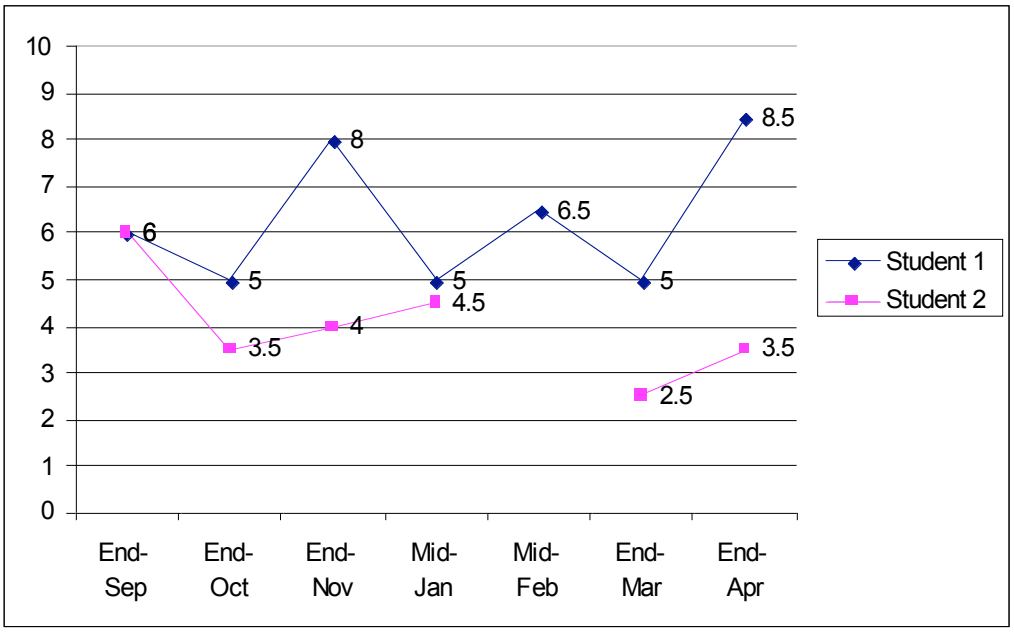


Figure 6: Individual Scores on Basic Skills Probes for each Special Education Algebra Student



Our second set of analyses relied on individual students' slope data, rather than the group means. As described earlier, each student's weekly growth rate was calculated using ordinary least square regression to determine his/her slope value. Table 14 presents the range, mean, and standard deviation of slopes for the Algebra Foundations and Content Analysis-Multiple Choice probes for all students in District A. Table 14 also indicates that on average students grew .39 and .54 points each week on the Algebra Foundations and the Content Analysis- Multiple Choice probes, respectively. This result suggests that Content Analysis-Multiple Choice probes may be slightly more sensitive in identifying student growth.

Table 14. Descriptive Statistics for Weekly Slope Values on Both Probes for General Education Classes in District A

	N	Range	Mean	SD
Algebra Foundations	76	-.27 – 1.37	.39	.24
Content Analysis-Multiple Choice	77	-.30 – 1.76	.54	.34

We were also interested in determining whether slopes on the three types of algebra progress monitoring measures differed by type of class. Table 15 reports the average slope values on each of the measures by class type. This table also included data for the Special Education Algebra students in District A and the yearlong algebra course in District C.

Table 15. Descriptive Statistics for Weekly Slope Values for Three Probes by Class Type

	N	Range	Mean	SD
Algebra Foundations				
PreAlgebra	15	-.07 – 1.37	.32	.32
Algebra 1	43	-.27 - .79	.35	.19
8th Grade Algebra	18	.25 - .90	.54	.18
Special Education Algebra	2	-.08 - .05	-.01	.10
Content Analysis-Multiple Choice				
PreAlgebra	16	.16 - .89	.37	.17
Algebra 1	43	-.30 – 1.14	.47	.31
8th Grade Algebra	18	.49 - 1.76	.87	.33
District C Yearlong Algebra	6	-.11 - .65	.43	.28
Basic Skills				
District C Yearlong Algebra	6	-.10 - .24	.02	.13
Special Education Algebra	2	-.03 - .22	.09	.18

As Table 15 reveals, students in 8th Grade Algebra demonstrated more growth on the Algebra Foundations and Content Analysis-Multiple Choice probes than did students in the other types of classes. Similarly, students in Algebra 1 had higher mean slopes than did students in PreAlgebra. For all general education classes, the Content Analysis-Multiple Choice measure produced higher mean slopes than either the Basic Skills or the Algebra Foundations measures. For students in Special Education Algebra, neither probe detected changes in student performance. The Basic Skills probe did not reflect student growth for either the District C yearlong algebra students or for students in the Special Education Algebra class (though readers should note that the respective class sizes were 6 and 2, which limits the generalizability of any findings for these groups).

It is important to note that we used a weekly growth rate of .5 as benchmark and goal in our research. We anticipate that in order for algebra progress monitoring measures to be useful to teachers on a practical level, they must be able to expect to see scores grow by at least one point every two weeks (hence a weekly growth rate of .5). The results of this study suggest that these measures are sufficiently sensitive for 8th Grade Algebra students for either measure, and for Algebra 1 students with the Content Analysis-Multiple Choice measure. The growth of PreAlgebra students was somewhat lower that our benchmark.

Finally, we were interested in determining whether the growth rates students obtained on the two types of probes were associated with other indicators of growth. To do this, we examined the correlations between the slope values of both types of probes and teacher growth ratings and IAAT gain scores. These results are presented in Table 16. We found that there was a small relationship between IAAT gain scores and slope values in both types of probes. We did not find correlation between teacher ratings and both types of probes. This finding did not surprise us because there was small variance in Teacher Growth Rating.

Table 16. Correlations Between Slope Values and Teacher Growth Ratings and IAAT Gain Scores

	All Students	
	Teacher Growth Rating	IAAT Gain Score
Algebra Foundations Slope	.03	.32**
Content Analysis-Multiple Choice Slope	.08	.22

** $p < .01$

Summary and Future Research

The main purpose of this study was to assess the reliability and validity of the Algebra Foundations and the Content-Analysis-Multiple Choice probes in an Iowa school district and to examine the utility of the measures for screening student progress. Seventy-eight students in grades eight to twelve enrolled in general education algebra classes in District A participated in the study. When examining student progress, we also included two students in District A enrolled in a special education algebra course and six students from District C for whom we had yearlong data. Data were gathered from September 2005 to April 2006. Over eight months of data collection, general education students in District A completed two Algebra Foundations probes and two Content Analysis-Multiple Choice probes each month. Special education algebra students in District A completed two Algebra Foundations probes and two Basic Skills probes each month. District C students in the yearlong algebra course completed two Basic skills probes and two Content Analysis-Multiple Choice probes each month. Given the limited sample sizes for the Special Education Algebra class and the District C yearlong algebra class, we have constrained our comments in this summary to the primary District A sample.

We assessed the alternate form reliability and test-retest reliability and criterion validity of the Algebra Foundations and Content Analysis-Multiple Choice probes. Our findings revealed that both types of probes possessed adequate levels of reliability with the exception of the alternate reliability of the first Content Analysis-Multiple Choice probe. We specifically hypothesized that as the semester went on and students became more familiar with the probes, both types of reliability would increase. Our findings partially supported this hypothesis for both probes types since there were fluctuations in the reliability scores over the year. However, it is important to note that reliability scores were higher at the end of the year than they were at the beginning of the year.

To assess the validity of both types of probes, we gathered data from a variety of indicators of students' proficiency in algebra including course grades, teachers' evaluations of student proficiency and growth, and performance on standardized assessment instruments including IAAT, ITED, and ITBS scores. We assessed two types of validity: concurrent and predictive.

We assessed concurrent validity by examining the relationship of probe scores to teachers' evaluation of their students' proficiency, IAAT, ITED, and ITBS scores. We found that probe scores were moderately or highly correlated with these criterion measures with the

exception that there was no correlation between ITBS scores and probe scores. When we disaggregated the data to examine concurrent validity by class type, we found widely differing coefficients by class type, with the criterion validity coefficients for both types of probes nonsignificant for 8th Grade Algebra students. Although the small sample sizes on which this finding is based dictate that we interpret it with caution, this finding does suggest that the criterion validity of the measures may vary for different levels of algebra and/or students with varying backgrounds in algebra.

We assessed predictive validity by investigating the relationship of the earliest probe scores (taken in the Mid-September and End-September) to teacher ratings of growth, students' end-of-term algebra grades, IAAT scores taken at the end of the course, ITED scores and ITBS scores. Our results revealed that in general, both probes were moderately correlated with IAAT and ITED scores, but not with teacher ratings of growth and ITBS. When examining the predictive validity of the probes using data disaggregated by class type, we again found remarkable differences among class types. For PreAlgebra students, both the Algebra Foundations and the Content Analysis-Multiple Choice probes had moderate or high correlations with Post IAAT and ITED scores, but not with teacher growth rating, and algebra grade. For Algebra 1 students, both types of measures were moderately correlated with Post IAAT, ITED scores, and Algebra Grades, but not with teacher growth ratings. For 8th Grade Algebra students, both types of probes had low or moderate correlations with teacher growth ratings, but not with other criterion measures. This result implies that the probes may function differently for students of varying algebra backgrounds and/or ability levels.

We were also interested in examining whether the Algebra Foundations and Content Analysis-Multiple Choice probes reflected growth similarly for each type of class. To address this issue, three sets of analyses were conducted. First, we examined graphs of the mean scores for students in each of the three class types for each data collection session. We found fluctuations in the Algebra Foundations probe scores for all the class types, with limited growth across the first three months. When examining the growth in the Content Analysis –Multiple Choice probes, we found more regular patterns of increases across the school year for all three class types. It is worth noting that the Content Analysis-Multiple Choice were more sensitive to changes in students performance than the Algebra Foundations probes. For all types of classes, mean scores in the last administration of the Content Analysis-Multiple Choice probes were almost twice as high as the mean scores in the first administration.

Next, we calculated individual student slope values for each type of probe and computed weekly rates of growth. Across all classes, we found average weekly growth rates of .39 on the Algebra Foundations measure and .54 on the Content Analysis-Multiple Choice measure. When we disaggregated the data by class type, only the 8th Grade Algebra students grew more than .5 units on either type of probe (though the mean growth rate for Algebra 1 students was .47 for the Content Analysis-Multiple Choice measure). Finally, we were also interested in to see if the growth that students showed on both types of probes was related to their teachers' ratings of growth and IAAT gain scores. We found that there was a small correlation between IAAT gain scores and slope values. No correlation existed between either type of probes and teacher ratings of growth.

APPENDIX

Standardized Administration Directions: *Algebra Foundations*

Standardized Administration Directions: *Content Analysis-Multiple Choice*

Standardized Administration Directions: *Basic Skills*

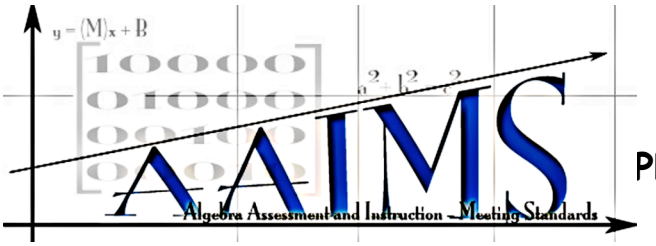
Algebra Foundations– Form 31

Content Analysis-Multiple Choice – Form 31

Basic Skills – Form 31

Teacher Rating of Student Proficiency

Teacher Rating of Student Growth



Project AAIMS 2005-06
XXX Senior High School

PROBE STANDARD DIRECTIONS
Algebra Foundations Probes

GENERAL INTRODUCTION TO PROGRESS MONITORING:

The **FIRST** time you administer algebra probes, say:

*As you know, your class and other algebra classes at XXX Senior High are working with Iowa State on a research project to learn more about improving algebra teaching and learning. Twice each month, we will be doing short algebra assessments, or probes, to monitor your learning in algebra. Remember that **all** students will be completing the probes and I will see the scores for all students, but your score will **only** be used for the Project AAIMS research if both you and your parent/guardian have given permission.*

There are a few things you should know about these probes. **First**, you will be given a limited amount of time to work on the problems. These probes are different from classroom tests or quizzes and are not meant to be completely finished. What's important is that as you learn more about algebra in this class, your scores will improve. **Second**, keep in mind that the object of the probe is to correctly answer as many questions as you can in the amount of time given. There may be problems on the probes that are difficult or unfamiliar. Please look at each problem. If you do not know how to answer it, skip it, and go on to the next problem. DO NOT spend a great deal of time on any one problem. If you get to the end of the probe and still have time to work, go back to the problems you skipped and try to solve them. **Third**, your scores on these probes will be used to see your progress in algebra. Because of this, it's important that you try your best. Do you have any questions at this point?

ALGEBRA FOUNDATIONS PROBES:

Hand out probes (B-31), keeping them face down. Ask students to keep the probes face down and write their name and the date on the back of the probe.

Give the standard directions:

The **FIRST** time you administer ALGEBRA FOUNDATIONS algebra probes, say:

The problems on this probe include translating words into expressions, solving simple equations, interpreting line graphs, and completing function or pattern tables. Please look at each problem and decide if you know how to do it. If you do, go ahead and solve the problem. If aren't certain or think you can't solve the problem, skip it and move to the next one. Don't spend too much time on any one problem. The object of the probe is to answer as many problems correctly in the time available. Once you get to the end, go back and work on the difficult problems. This probe starts with a sample page so you have a chance to practice doing these kinds of problems.

Please turn your paper over. This sample page shows some examples of the types of problems on the Algebra Foundations probes. As you can see, there are several different kinds of problems. You are to fill in the blanks, empty boxes or write an expression or word phrase for each problem. It is okay to skip around on this probe and work on the problems you think are the easiest first.

Let's take a minute so you can practice doing an Algebra Foundations probe. If you finish before I say 'Stop', please do NOT turn to the next page. Any questions? Ready, begin. *[time for 1 minute]* Stop, pencils down.

Now that you've had a chance to try out this type of probe, do you have any questions? *[Only answer procedural questions—do not suggest ways to solve the problems.]*

Now we'll do the first Algebra Foundations probe. You will have 5 minutes to work on this two-page probe. Remember, your job is to answer as many problems correctly as you can in 5 minutes. Please look at each problem. If you do not know how to do it, skip it and move on. If you get to the end of the probe before the time is up, go back and work on the problems you skipped.

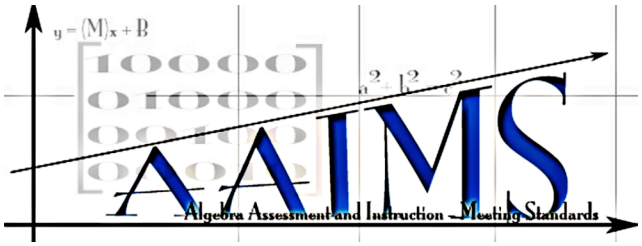
When I say begin, please turn past the sample page and begin working. You will have 5 minutes. Please do your best work. Ready? Begin

Time for 5 minutes. When the timer goes off, say Stop. Please put your pencils down, *and collect student papers.*

*For **ALL OTHER** administrations, hand out the probes face down and say*

Please write your name, my name and the date on the back of your paper. You are going to do an Algebra Foundations probe. You will have 5 minutes to work. Remember to try and complete as many problems correctly as you can in the time allowed. If you're not sure of an answer, skip the problem and come back to it if you have time left. When you are simplifying, be sure to write your answer in lowest terms. Please do your best work. Ready? Begin.

Time for 5 minutes. When the timer goes off, say Stop. Please put your pencils down, *and collect student papers.*



PROBE STANDARD DIRECTIONS
Textbook Probes

E (TEXTBOOK) PROBES:

Hand out probe E-31 (with the sample page), keeping them face down. Ask students to keep the probes face down and write their name and the date on the back of the probe.

Give the standard directions:

The **FIRST** time you administer the TEXTBOOK algebra probes, say:

The problems on this probe come from the chapters of the book, but they are not in any special order. For example, a problem from Chapter 1 could be the last problem on the probe. Please look at each problem and decide if you know how to do it. If you do, go ahead and solve the problem. If aren't certain or think you can't solve the problem, skip it and move to the next one. Don't spend too much time on any one problem. The object of the probe is to answer as many problems correctly in the time available. Once you get to the end, go back and work on the difficult problems. Remember that you may earn partial credit by showing your work even if you can't solve the entire probe. Do NOT make wild guesses because this will cause you to lose points on the probe.

Please turn your paper over. This sample page shows some examples of the types of problems on the Textbook probes. The problems on this probe are drawn from the different types of problems you are learning in the textbook. The questions are multiple choice. Each problem is worth 3 points, but you can earn partial credit by showing your work. Unless you are completely certain of the correct answer, the best strategy is to show your work. If you do not know the answer, you should NOT make wild guesses. You will lose points from your total score on the probe when you make wild guesses.

Look at the three boxes in the first row labeled A, B, and C. You'll notice that all three have answers and that the problem is the same for all three. Look at the box for Student A. She thought she knew the correct answer, so she just circled her choice at the bottom. Unfortunately, she was incorrect, so she will lose

a point for this problem. Student B showed his work, but did not know how to finish the problem. Because he did part of the problem correctly, Student B will earn 1 out of 3 points on this problem. Student C solved the problem, but made an error, so her final answer is not correct. Because she showed her work, she will earn 1 out of 3 points on the problem for the part she has done correctly. As you can see from these examples, it is important to show your work on these probes.

Let's take a minute so you can practice doing a Textbook probe. If you finish before I say 'Stop', please do NOT turn to the next page. Any questions? Ready, begin. [Time for 1 minute] *Stop, pencils down.*

Now that you've had a chance to try out this type of probe, do you have any questions? [Only answer procedural questions—do not suggest ways to solve the problems.]

Now we'll do the first Textbook probe. You will have 7 minutes to work on this two-page probe. Remember, your job is to answer as many problems correctly as you can in 7 minutes. Please look at each problem. If you do not know how to do it, skip it and move on. If you get to the end of the probe before the time is up, go back and work on the problems you skipped. Remember that you may earn partial credit by showing your work even if you can't solve the entire problem. Do NOT make wild guesses because this will cause you to lose points on the probe.

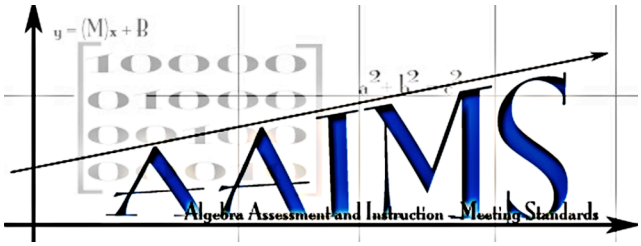
When I say begin, please turn past the sample page and begin working. You will have 7 minutes. Please do your best work. Ready? Begin.

Time for 7 minutes. When the timer goes off, say *Stop. Please put your pencils down*, and collect student papers.

For **ALL OTHER** administrations, hand out the probes face down and say

Please write your name and the date on the back of your paper. You are going to do a Textbook probe. You will have 7 minutes to work. Remember to try and complete as many problems correctly as you can in the time allowed. If you're not sure of an answer, skip the problem and come back to it if you have time left. Remember that you can earn partial credit by showing your work. Do NOT make wild guesses. Please do your best work. Ready? Begin.

Time for 7 minutes. When the timer goes off, say *Stop. Please put your pencils down,* and collect student papers.



Project AAIMS 2005-06
XXX High School

PROBE STANDARD DIRECTIONS
Basic Skills Probes

GENERAL INTRODUCTION TO PROGRESS MONITORING:

The **FIRST** time you administer algebra probes, say:

*As you know, your class and other algebra classes at XXX High are working with Iowa State on a research project to learn more about improving algebra teaching and learning. Twice each month, we will be doing short algebra assessments, or probes, to monitor your learning in algebra. Remember that **all** students will be completing the probes and I will see the scores for all students, but your score will **only** be used for the Project AAIMS research if both you and your parent/guardian have given permission.*

There are a few things you should know about these probes. First, you will be given a limited amount of time to work on the problems. These probes are different from classroom tests or quizzes and are not meant to be completely finished. What's important is that as you learn more about algebra in this class, your scores will improve. Second, keep in mind that the object of the probe is to correctly answer as many questions as you can in the amount of time given. There may be problems on the probes that are difficult or unfamiliar. Please look at each problem. If you do not know how to answer it, skip it, and go on to the next problem. DO NOT spend a great deal of time on any one problem. If you get to the end of the probe and still have time to work, go back to the problems you skipped and try to solve them. Third, your scores on these probes will be used to see your progress in algebra. Because of this, it's important that you try your best. Do you have any questions at this point?

BASIC SKILLS PROBES:

Hand out probe A-31 (with the sample page), keeping the probes face down. Ask students to keep the probes face down and write their name and the date on the back of the probe.

Give the standard directions:

The **FIRST** time you administer BASIC SKILLS algebra probes, say:

Please turn your paper over. This sample page shows some examples of the types of problems on the Basic Skills probes. The questions include solving algebra equations using basic math facts, simplifying expressions by combining like terms, using the distributive property to simplify expressions, and solving proportion, or ratio problems. Now we'll take a minute so <u>you</u> can practice doing a Basic Skills probe. If you finish before I say 'Stop', please do NOT turn to the next page. Any questions? Ready, begin. [Time for 1 minute] *Stop, pencils down.*

Now that you've had a chance to try out this type of probe, do you have any questions? [Only answer procedural questions—do not suggest ways to solve the problems.]

Now we'll do the first Basic Skills probe. You will have 5 minutes to work on this two-page probe. Remember, your job is to answer as many problems correctly as you can in 5 minutes. Please look at each problem, but if you do not know how to do it, skip it and move on. If you get to the end of the probe before the time is up, go back and work on the more difficult problems. When you solve the simplifying questions, be sure to go as far as you can with your answer.

When I say begin, please turn past the sample page and begin working. You will have 5 minutes. Please do your best work

Time for 5 minutes. When the timer goes off, say *Stop. Please put your pencils down*, and collect student papers.

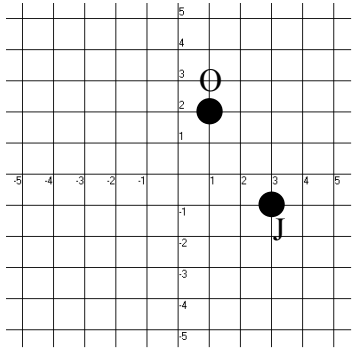
For **ALL OTHER** administrations, hand out the probes face down and say

Please write your name and the date on the back of your paper. You are going to do a Basic Skills probe. You will have 5 minutes to work. Remember to try and complete as many problems correctly as you can in the time allowed. When you are simplifying, be sure to go as far as you can with your answer [write your answer in lowest terms]. Please do your best work. Ready? Begin.

Time for 5 minutes. When the timer goes off, say *Stop. Please put your pencils down*, and collect student papers.

Find the ordered pair for each point:

J(,) O(,)



Fill in the empty box:

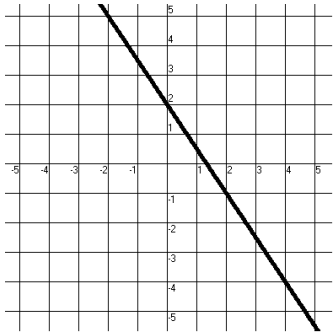
s	$3s$
6	18
7	21
8	
9	27

Fill in the empty box:

n	$4n + 7$
-1	3
-2	-1
-3	
-4	-9

Fill in the empty box:

b	
5	2
3	0
0	-3
-2	-5



What is the slope?

What is the y intercept?

If $y > 9$, two possible values for y are _____ and _____

Evaluate:
$9 \bullet 4 - 6$

Simplify:
$7f + (2f + f)$

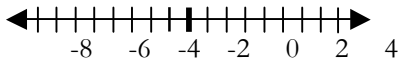
Solve:
$n + 3 = 8$
$n =$

Evaluate $4b + 2$ when
$b = 1$ _____ and when
$b = 3$ _____

Write the expression for this phrase:
6 less than a number

Evaluate:
$(-2) \bullet (-4)$

Graph the expression $m > 6$



Write a word phrase for this expression:
$n + 9$

Evaluate:
$4 + (9 \div 3) - 2^2$

Evaluate:
$(-2)^3$

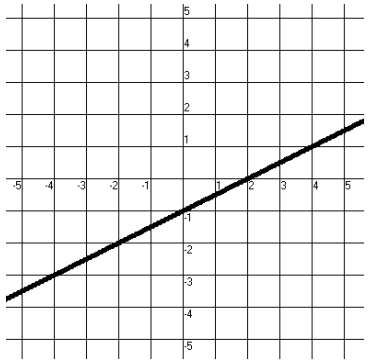
Write the expression for this phrase:
9 multiplied by a number

Evaluate $2x + y$ when
$x = 2$ and $y = 3$

Write a word phrase for this expression:
$10b - 7$

If $2a + 4 < 20$, two possible values for a are _____ and _____

Simplify:
$6 - 2(b - 4)$



What is the slope?

What is the *y* intercept?

Fill in the empty box:

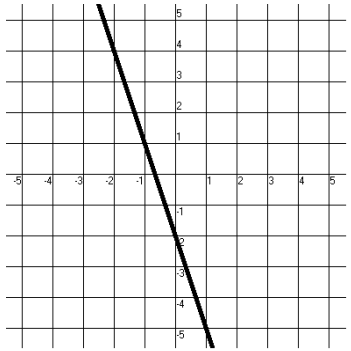
n	
6	4
9	6
12	8
15	10

Fill in the empty box:

t	$t - 7$
-2	-9
2	-5
6	
10	3

Fill in the empty box:

w	
4	11
6	17
8	23
10	29



What is the slope?

What is the *y* intercept?

Write a word phrase for this expression:
$x \div 4$

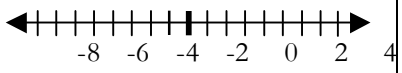
Evaluate:
$(-16) \div (-4)$

Write the expression for this phrase:
8 more than twice a number

Solve:
$15 - 8 = x$
$x =$

Evaluate
4^2

Graph the expression $p \leq -3$



Simplify:
$9x + 3 - 4(x + 9)$

Write the expression for this phrase:
10 divided by a number

Evaluate $8g - 4$ when
$g = 2$ _____
$g = 4$ _____

Solve:
$24 \div x = 6$
$x =$

Evaluate:
$10 - 3 \bullet 5 + 2$

Simplify:
$12n - 5 + 3 - 7n$

Write a word phrase for this expression:
$h \bullet 5$

Evaluate:
$(-3)(9 - 7)$

Evaluate:
$\sqrt{81}$

Solve:
$6t = 36$
$t =$

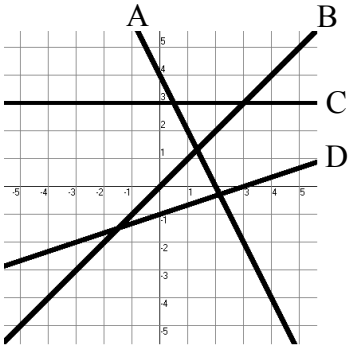
Solve:
$3x + 4 = 19$

a) $x = 8$
b) $x = 22$
c) $x = 15$
d) $x = 5$

Evaluate $a^2 - b \div 2$ when $a = 4$ and $b = 6$

a) 1
b) 5
c) 10
d) 13

Which line on the graph is $y + 2x = 4$?



a) Line A
b) Line B
c) Line C
d) Line D

Simplify:
$3(m + 2) + 2(m - 1)$

a) $5m + 4$
b) $5m + 1$
c) $6m + 8$
d) $6m - 8$

Evaluate the expression:

4^{-2}

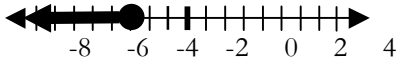
a) -16
b) $\frac{1}{16}$
c) $\frac{1}{8}$
d) -8

Solve the linear system:

$x - y = 4$
$x + 2y = 19$

a) $(-1, -5)$
b) $(5, 8)$
c) $(-2, 19)$
d) $(9, 5)$

This graph shows the solution for which equation?



a) $x > -3$
b) $2x \leq -6$
c) $-3x > 9$
d) $3x \geq 9$

Write the equation in slope-intercept form if $m = \frac{1}{2}$ and $b = 3$

a) $y = 2x + 3$
b) $y = 3x + \frac{1}{2}$
c) $x = \frac{1}{2}y - 3$
d) $y = \frac{1}{2}x + 3$

Evaluate $d + 3e^2$ when $d = 5$ and $e = 2$ a) 11 b) 23 c) 17 d) 10	Solve: $6c + 4 = -3c - 14$ a) $-\frac{10}{3}$ b) -2 c) 2 d) 6	Find the slope of a line through $(1, -1)$ $(5, 2)$ a) $\frac{1}{5}$ b) $\frac{3}{4}$ c) -6 d) $-\frac{4}{3}$	Simplify: $6(2b - 3) - 3(2 - b)$ a) $15b - 24$ b) $9b - 9$ c) $9b + 12$ d) $15b + 12$
Simplify the expression: $\frac{a^2}{ab^3} \bullet \frac{b^4}{a^3}$ a) $\frac{a^8}{a^3b^3}$ b) $\frac{ab^8}{a^4b^3}$ c) $\frac{b}{a^2}$ d) $\frac{b}{a}$	Solve the linear system: $-6x + 3y = -6$ $2x + 6y = 30$ a) (6, 3) b) (3, 4) c) (2, 6) d) (4, – 3)	Simplify $b^2 - 4b + 2b^2 + 7 - 5$ a) $3b^2 - 4b + 2$ b) $2b + 2$ c) $-b^2 - 4b + 12$ d) $3b^2 - 4b + 12$	Write the equation of a line through (5, 3) (4, 9). Use point-slope form. a) $y + 1 = 2(x - 4)$ b) $y + 4 = -6(x - 1)$ c) $y - 3 = -6(x - 5)$ d) $y = -6x + 30$

Solve: $9 + a = 15$ $a =$	Solve: $10 - 6 = g$ $g =$
Evaluate: $12 + (-8) + 3$	Simplify: $9 - 4d + 2 + 7d$
Simplify: $2x + 4 + 3x + 5$	Simplify: $5(b - 3) - b$
Solve: $12 - e = 4$ $e =$	Solve: $q \bullet 5 = 30$ $q =$
Simplify: $4(3 + s) - 7$	Evaluate: $8 - (-6) - 4$
Simplify: $b + b + 2b$	Simplify: $2 + w(w - 5)$
Solve: $\frac{b}{6} = \frac{12}{18}$ $b =$	Solve: 1 foot =12 inches 5 feet = ____ inches
Simplify: $7 - 3(f - 2)$	Simplify: $4 - 7b + 5(b - 1)$
Evaluate: $-5 + (-4) - 1$	Simplify: $s + 2s - 4s$
Solve: $63 \div c = 9$ $c =$	Solve: $x + 4 = 7$ $x =$
Simplify: $2(s - 1) + 4 + 5s$	Simplify: $-5(q + 3) + 9$
Simplify: $8m - 9(m + 2)$	Evaluate: $9 + (-3) - 8$
Solve: 3 feet = 1 yard ____ feet = 9 yards	Solve: $\frac{12}{2} = \frac{48}{e}$ $e =$
Evaluate: $4 - (-2) + 8$	Simplify: $y^2 + y - 4y + 3y^2$
Simplify: $2k + 3 - 5(k + 7)$	Simplify: $3(c + 2) - 2c$

Solve: $3 \bullet 8 = m$ $m =$	Solve: $\frac{1.5}{3} = \frac{h}{9}$ $h =$
Evaluate: $-9 + 5 + 8$	Simplify: $7b - 4 - 3 - 2b$
Simplify: $x + 2(x - 5) - 3$	Simplify: $2e - 3(e - 4)$
Solve: $d - 5 = 4$ $d =$	Solve: $6 + 7 = v$ $v =$
Simplify: $5(3 + f) - 2f + 6$	Evaluate: $-5 + 6 - 6$
Simplify: $5 - 2b + 4(b + 3)$	Simplify: $4 + 10(1 - r)$
Solve: 4 quart = 1 gallon ____ quarts = 3 ¼ gallons	Solve: 2.5 cm = 1 inch ____ cm = 6 inches
Simplify: $4(y + 1) - 8y$	Simplify: $6a + 2a - 9 + 3a^2$
Evaluate: $14 - 7 + (-3)$	Evaluate: $-1 + 4 + (-7)$
Solve: $\frac{36}{6} = s$ $s =$	Solve: $\frac{500}{j} = \frac{10}{2}$ $j =$
Simplify: $-3w^2 + 5w^2 - 5 + 12$	Simplify: $-3(u + 3) - 2u + 5$
Simplify: $9 - 4(v + 2)$	Simplify: $2c - 3c - c$
Solve: $4r = 28$ $r =$	Solve: $h \div 6 = 8$ $h =$
Simplify: $16 + 2(t - 4) - 3t$	Evaluate: $-2 + (-5) + (-8)$
Simplify: $c - 3(c + 2) + 8$	Simplify: $3z - 8z + 2 + 9$

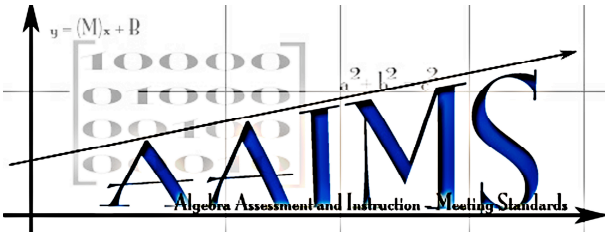
Teacher ____________________

Project AAIMS:
Algebra Assessment and Instruction:
Meeting Standards
XXX High School
Fall 2005

<u>Directions</u>:	For each student, rate his or her **general proficiency in algebra** relative to other students in your class(es). A rating of "1" indicates a very low level of proficiency, "4" indicates average proficiency, and "7" indicates exceptional proficiency. Try to spread student ratings across the **full range of the scale**, not clustering students only in the middle or toward one end.

<u>Student</u>	<u>Algebra Proficiency</u>						
	Low			Average			High
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7
	1	2	3	4	5	6	7

Teacher _____



Project AAIMS:
Algebra Assessment and Instruction:
Meeting Standards
XXX Senior High

Teacher Rating of Student Progress

Directions: Below is a list of the students you teach. Please rate the amount of progress or improvement each student has made in algebra during this course. A rating of "1" indicates no growth or a decrease in level of performance, "3" indicates average progress, and "5" indicates exceptional progress, far beyond what you expected. If the list includes a student who has dropped the class, just draw a line through his/her name. Thank you!

Student	Algebra Progress				
	Low		Average		High
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5